

Theory of Kondo suppression of spin polarization in nonlocal spin valves

K.-W. Kim,^{1,2,3,*} L. O'Brien,^{4,5,6,*} P. A. Crowell,⁷ C. Leighton,^{8,†} and M. D. Stiles^{1,‡}

¹Center for Nanoscale Science and Technology, National Institute of Standards and Technology, Gaithersburg, Maryland 20899, USA

²Maryland NanoCenter, University of Maryland, College Park, Maryland 20742, USA

³Institut für Physik, Johannes Gutenberg Universität Mainz, Mainz 55128, Germany

⁴Department of Chemical Engineering and Materials Science, University of Minnesota, Minnesota, 55455 USA

⁵Thin Film Magnetism, Cavendish Laboratory, University of Cambridge, CB3 0HE, UK

⁶Department of Physics, University of Liverpool, Liverpool L69 7ZE, UK

⁷School of Physics and Astronomy, University of Minnesota, Minnesota 55455, USA

⁸Department of Chemical Engineering and Materials Science, University of Minnesota, Minnesota 55455, USA

(Dated: February 10, 2017)

We theoretically analyze contributions from the Kondo effect to the spin polarization and spin diffusion length in all-metal nonlocal spin valves. Interdiffusion of ferromagnetic atoms into the normal metal layer creates a region in which Kondo physics plays a significant role, giving discrepancies between experiment and existing theory. We start from a simple model and construct a modified spin drift-diffusion equation which clearly demonstrates how the Kondo physics not only suppresses the electrical conductivity but even more strongly reduces the spin diffusion length. We also present an explicit expression for the suppression of spin polarization due to Kondo physics in an illustrative regime. We compare this theory to previous experimental data to extract an estimate of the Elliot-Yafet probability for Kondo spin flip scattering of 0.7 ± 0.4 , in good agreement with the value of $2/3$ derived in the original theory of Kondo.

I. INTRODUCTION

Pure spin currents, devoid of charge current flow, are now routinely generated in metals-based systems [1] via a number of techniques, including the use of thermal gradients [2], the spin Hall effect [3], spin pumping [4], and nonlocal spin injection [5], each method providing a unique insight into spin relaxation. In particular, the ability to separate charge and spin currents using the nonlocal spin valve [5, 6], thereby circumventing difficulties interpreting ‘local’ spin valve measurements, makes it one of the most unambiguous techniques for probing spin transport. This geometry is especially useful at the nanoscale, where isolating the factors affecting spin accumulation, diffusion, and relaxation, both within the bulk and across interfaces, represents a pressing problem [7–15]. Indeed, examining the role of specific defects in relaxing spins in metals at this length scale – including interfaces, grain boundaries, and magnetic and highly spin-orbit coupled impurities – will be critical for realizing future low resistance-area-product spintronic devices, e.g. current perpendicular-to-plane giant magnetoresistance sensors [16].

A nonlocal spin valve consisting of a normal metal channel connected by two ferromagnetic contacts is illustrated in Fig. 1(a): the injected current I_{21} generates a spin accumulation at the interface between the nonmagnet and ferromagnet (Lead 2). This accumulation diffuses in both directions down the channel causing a pure spin current to flow towards Lead 1, which decays on a characteristic spin diffusion length, l_N^{sf} . The remaining spin population reaching Lead 3 generates a nonlocal voltage difference V_{34} between the ferromag-

netic contact and channel and therefore a nonlocal resistance, $R_{\text{NL}} = \Delta V_{34}/I_{21}$. The sign of this resistance depends on the relative orientation of the two ferromagnets, and so by applying a magnetic field to alternate the ferromagnetic contact magnetization from parallel to antiparallel, a nonlocal spin signal, ΔR_{NL} , is measured, directly related to the magnitude of the spin accumulation under the contact.

In relatively simple all-metal nonlocal spin valves (e.g., $\text{Ni}_{80}\text{Fe}_{20}/\text{Cu}$) that are fabricated from nominally high-purity materials, the standard theory of spin drift-diffusion developed by Valet and Fert [17], combined with the Elliott-Yafet spin relaxation mechanism [18–20], which predominates in light metals, dictates that l_N^{sf} , the spin accumulation, and therefore ΔR_{NL} should monotonically increase as temperature T decreases. Surprisingly, however, ΔR_{NL} is widely found to anomalously *decrease* at low T in $\text{Ni}_{80}\text{Fe}_{20}/\text{Cu}$, Fe/Cu and Co/Cu nonlocal spin valves [21–30], even when the resistivity of the normal metal and the ferromagnet, $\rho_N(T)$ and $\rho_F(T)$, are found to continuously decrease on cooling.

Consensus is emerging that this unexpected reduction of ΔR_{NL} at low T is due to spin relaxation at dilute magnetic impurities [26–28], with recent results demonstrating that a manifestation of the Kondo effect is at the heart of the suppression [29, 30]. The Kondo effect [31] arises in metals with dilute magnetic impurities, as a result of s - d exchange between the conduction electrons and virtual bound impurity states. This exchange results in an additional higher order contribution to the scattering cross section, proportional to $\log T$, which can dominate in otherwise highly pure metals at low T . In charge transport, the classic signature of the Kondo effect is an increase in the conduction electron scattering rate at low T , resulting in a minimum in resistance (maximum in conductance) and a logarithmic increase in $\rho(T)$ about a characteristic temperature T_K [32]. Similarly, for spin transport the additional (spin-flip) scattering was recently found to efficiently relax the spin accumulation, suppressing ΔR_{NL} with

*K.-W. Kim and L. O'Brien contributed equally to this work.

[†]Electronic address: leighton@umn.edu

[‡]Electronic address: mark.stiles@nist.gov

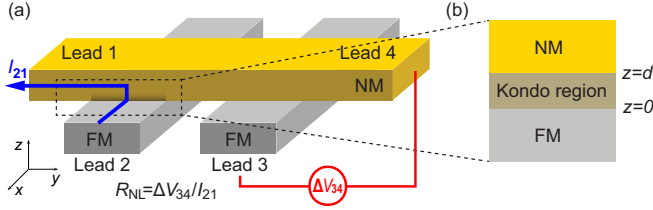


FIG. 1: (color online) (a) Geometry of a nonlocal spin valve, consisting of a normal metal (NM) channel and two ferromagnetic (FM) contacts. When an electrical current I_{21} is applied (blue line), a nonequilibrium spin accumulation develops, giving rise to a finite voltage difference ΔV_{34} and nonlocal resistance $R_{NL} = \Delta V_{34}/I_{21}$. (b) Schematic of the model used to determine the suppression of spin polarization at a ferromagnet/nonmagnet interface ($z = 0$) due to interdiffused ferromagnetic atoms over a characteristic length scale d , which we term the Kondo region.

what is also observed to be a $\log T$ dependence [29]. This occurs even for nonmagnetic channels that are largely impurity-free throughout the bulk, due to inevitable interdiffusion at the ferromagnet/nonmagnet interface. The situation is schematically depicted in Fig. 1(b), where interdiffusion creates a region with ‘high’ levels of ferromagnetic impurities (on average ≈ 100 ’s $\mu\text{mol/mol}$ [47]) which rapidly relax the injected spins at the interface, reducing the effective polarization of the bias current [29]. To this point, a direct quantitative link has already been established between the degree of interdiffusion and magnitude of Kondo suppression in nonlocal spin valves [30]. Reciprocally, the disruption of the Kondo singlet through the injection of sufficiently large spin currents has also been investigated [33, 34]. Since there are no magnetic impurities far away from the interface, spin diffusion in the bulk is described by the spin drift-diffusion equation in the Valet-Fert theory, and a measure of $\rho_N(T)$ yields no indication of the Kondo effect. Naturally, in devices where impurity levels are sufficiently high throughout the nonmagnet, either due to intentional doping [28], source contamination [27], or contamination during deposition [26], the effects of Kondo scattering can equally be found to enhance spin relaxation in the bulk of the channel, thereby reducing τ_N^{sf} . Despite this growing body of experimental work, a complete theoretical treatment of the effect remains outstanding; a description of the suppression of the spin polarization near the interface due to the Kondo effect is therefore the aim of this work.

In this paper, we start from the Boltzmann equation and follow the Valet-Fert theory [17] to construct a modified spin drift-diffusion equation which is valid in the presence of dilute magnetic impurities. The Valet-Fert theory is an approximation to the Boltzmann equation, neglecting variations of the distribution function on the length scale of the mean free path. Thus, it is a valid approximation for structures with interfaces separated by more than a mean free path, which is the case here. While the Valet-Fert theory is developed for $T = 0$, we allow for nonzero temperature to extract Kondo contributions. Then, we project our theory to a low-temperature regime, keeping the additional Kondo contributions. Using the modified spin drift-diffusion equation, we compare our

theory to experimental data to extract an estimate of the Elliot-Yafet parameter for Kondo spin relaxation. This is found to be in very good agreement with the value originally proposed by Kondo [31] as well as resolving the discrepancy between the standard Valet-Fert theory and experiment [30].

The paper is organized as follows. In Sec. II, we develop a theory describing suppression of the spin polarization at the interface. We first present the theory without derivation in Sec. II A and then we demonstrate that the spin polarization at the interface has a maximum at a finite temperature. In Sec. III, we compare the theory to our experimental data [30]. In Sec. IV, we present mathematical details which are referred to in Sec. II. Finally, in Sec. V we summarize the paper.

II. INTERFACIAL KONDO EFFECT

A. Modification of the electrical conductivity and the spin diffusion length due to the Kondo effect

Starting from antiferromagnetic exchange coupling between conduction electrons and dilute magnetic impurities, Kondo [31] showed that electrical conductivity in metals is suppressed at low temperature. This is equivalent to suppression of the momentum relaxation time:

$$\frac{1}{\tilde{\tau}_N} = \frac{1}{\tau_N} + \frac{1}{\tau_K^{\text{eff}}}, \quad (1)$$

where τ_N is the momentum relaxation time without dilute magnetic impurities in the normal metal at the Fermi level and $\tilde{\tau}_N$ is the modified momentum relaxation time in the presence of the Kondo effect. τ_K^{eff} is the effective Kondo relaxation time, for which the explicit expression is given below in Eq. (29). τ_K^{eff} has a logarithmic temperature dependence, and thus it can be comparable to or even dominate τ_N at low temperature for very low impurity concentrations.

In addition to suppressing the scattering time (increasing the scattering rate), as in Eq. (1), the dilute magnetic impurities also suppress the spin relaxation time τ_N^{sf} .

$$\frac{1}{\tilde{\tau}_N^{\text{sf}}} = \frac{1}{\tau_N^{\text{sf}}} + \frac{\eta}{\tau_K^{\text{eff}}}, \quad (2)$$

where $\tilde{\tau}_N^{\text{sf}}$ is the modified spin relaxation time due to the Kondo effect. Here, η is the spin-flip probability during each Kondo scattering event. The proportionality between the change in the momentum relaxation rate ($1/\tau_K^{\text{eff}}$) and the spin relaxation rate (η/τ_K^{eff}) is similar to that found for the Elliot-Yafet scattering mechanism. For Elliot-Yafet scattering, the contribution to the spin-flip scattering rate is given by $1/\beta\tau$ where $1/\tau$ is the contribution to the momentum scattering rate and β is the Elliot-Yafet parameter. Thus, η is the inverse of the Elliot-Yafet parameter for spin relaxation from Kondo impurities. The value of η is determined by the geometry of the Fermi surface. For the spherical Fermi surfaces that we consider here, $\eta = 2/3$, as shown in Sec. IV. We note that in Ref. [28], the spin-flip probability is claimed to be around 0.3 based on a semiclassical argument [35]. Strictly, however, the

semiclassical argument does not apply for the higher order interactions giving rise to the Kondo physics.

One remark on our notation is in order. Generally speaking τ_N and τ_N^{sf} are k -dependent functions. But here, we drop the k dependence and take the values only at the Fermi level for simplicity. In Sec. IV, we restore the k dependence for the derivation. In particular, $\tau_N(k_N^F)$ and $\tau_N^{\text{sf}}(k_N^F)$ in Sec. IV are respectively τ_N and τ_N^{sf} here, where k_N^F is the Fermi wavevector in the normal metal.

The changes in the relaxation times above imply changes in the electrical conductivity and the spin diffusion length. In Sec. IV, we show that the Valet-Fert theory for the spin drift-diffusion equation still holds even in the presence of dilute magnetic impurities, once we impose Eqs. (1) and (2). The Valet-Fert theory provides links between quantities in the Boltzmann equation (such as relaxation times) and quantities in the drift-diffusion equation (such as the electrical conductivity and spin diffusion length) as given below in Eqs. (19) and (20). Up to first order in the Kondo rate, the modified conductivity and spin diffusion length are

$$\tilde{\sigma}_N = \sigma_N \left(1 - \frac{\tau_N}{\tau_K^{\text{eff}}} \right), \quad (3)$$

$$\tilde{l}_N^2 = l_N^2 \left(1 - \frac{\tau_N + \eta \tau_N^{\text{sf}}}{\tau_K^{\text{eff}}} \right). \quad (4)$$

Equation (3) corresponds to Kondo's original work, i.e. the conductivity reduction due to the Kondo effect. Equation (4) is the spin counterpart of the original Kondo effect, a central aspect of this paper.

At this point it is worth noting that the Kondo effect can affect the spin diffusion length much more dramatically than it does the conductivity, since $(\tau_N^{\text{sf}})^{-1}$ is usually much smaller than τ_N^{-1} . For example, $\tau_N^{\text{sf}}/\tau_N \approx 10^3$ in Ref. [1]. Thus it is possible that $1 - \tilde{l}_N^2/l_N^2$ is noticeable even though $1 - \tilde{\sigma}_N/\sigma_N$ is negligible.

B. Suppression of the spin polarization at the interface

We now solve the spin drift-diffusion equation with the modified quantities in Eqs. (3) and (4) at the interface. Our model is illustrated in Fig. 1(b). We consider a ferromagnet ($z < 0$)/nonmagnet ($z > 0$) interface at $z = 0$. Near the interface, interdiffused ferromagnetic atoms create a region in which dilute magnetic impurities are present. For illustration, we assume that the impurity concentration is constant over $0 < z < d$ and suddenly drops to zero at $z = d$. We term the region $0 < z < d$ the Kondo region.

The spin drift-diffusion equation is given by the set of equations below [17].

$$\frac{e}{\sigma_s(z)} \partial_z j_s(z) = \frac{\mu_s(z) - \mu_{-s}(z)}{l_s^2(z)}, \quad (5)$$

$$\partial_z \mu_s(z) = \frac{e}{\sigma_s(z)} j_s(z), \quad (6)$$

where $s = \pm$ denotes the spin majority and minority bands, e is the electron charge, $j_s(z)$, $\mu_s(z)$, $\sigma_s(z)$, and $l_s(z)$ are respectively the current expectation value, the chemical potential, the electrical conductivity, and the spin diffusion length of the spin s band at position z . σ_s and l_s are parameters treated as position independent in most cases. We explicitly retain the position dependence to emphasize the dependence of the parameters on the regions; $z < 0$, $0 < z < d$, and $z > d$. More explicitly, the parameters in each region are given by

$$(\sigma_s(z), l_s(z)) = \begin{cases} (\sigma_{s,F}, l_{s,F}) & \text{for } z < 0, \\ (\tilde{\sigma}_N, \tilde{l}_N) & \text{for } 0 < z < d, \\ (\sigma_N, l_N) & \text{for } z > d. \end{cases} \quad (7)$$

Here the subscripts N and F refer to the normal metal and the ferromagnetic metal and the tildes refer to the Kondo region [48]. Since the physical parameters in the normal metal do not have spin dependence, we drop the subscript s for $z > 0$.

The general solutions of Eqs. (5) and (6) are obtained in Ref. [17]. The spin accumulation $\mu_+ - \mu_-$ decays exponentially over the effective spin diffusion length defined by $(l_F^{\text{sf}})^{-2} = l_{+,F}^{-2} + l_{-,F}^{-2}$ for $z < 0$, $(\tilde{l}_N^{\text{sf}})^{-2} = 2\tilde{l}_N^{-2}$ for $0 < z < d$, and $(l_N^{\text{sf}})^{-2} = 2l_N^{-2}$ for $z > d$. To match the experimental situation of Refs. [29] and [30], we assume transparent interfaces at $z = 0$ and $z = d$ where the spin chemical potential and the currents are continuous. The solutions of Eqs. (5) and (6) in the given situation are obtained in Appendix A. Here, we simply present spatial profiles of the spin accumulation $\mu_+(z) - \mu_-(z)$ and the spin current $j_+(z) - j_-(z)$ for a set of parameters in Fig. 2. For clear presentation, we scale the quantities. In this case, $\mu_+(z) - \mu_-(z)$ is normalized to its $z = 0$ value, and $j_+(z) - j_-(z)$ to $j_{\text{app}}/2.5$. Here $j_{\text{app}} = j_+(z) + j_-(z)$ is the applied charge current which is independent of position due to the conservation of electrical charge. The factor of 2.5 is simply to allow both quantities to be plotted on the same scale.

Figure 2 clearly shows that there is suppression of the spin polarization due to the Kondo region. Since there are no magnetic impurities for $z > d$, relaxation rates for $z > d$ are the same with and without the Kondo region. Thus, the spin accumulation calculated at $z = d$ indicates the suppression of ΔR_{NL} due to the Kondo effect. To quantify this suppression, we analytically evaluate the ratio of the accumulation at d in the presence and absence of a finite impurity concentration in the Kondo region (i.e. in the limit $\tau_K^{\text{eff}} \rightarrow \infty$). Here we compute the following expression up to $O((\tau_K^{\text{eff}})^{-1})$:

$$\kappa \equiv \frac{\mu_+(d) - \mu_-(d)}{\lim_{\tau_K^{\text{eff}} \rightarrow \infty} [\mu_+(d) - \mu_-(d)]} \equiv \frac{\alpha(d)}{\lim_{\tau_K^{\text{eff}} \rightarrow \infty} [\alpha(d)]}, \quad (8)$$

where $\alpha(z)$ is defined as the current polarization, $\alpha(z) = [j_+(z) - j_-(z)]/j_{\text{app}}$ [49]. At this point it is worth emphasizing that, since it is a quantitative indication of the strength of Kondo suppression, the suppression ratio κ represents a key parameter of this work. Furthermore, as α is directly measurable, κ uniquely represents an experimentally accessible spin transport parameter with which to compare theory and measurement. In evaluating Eq. (8) we retain only terms up to $O(d)$, assuming that d is much shorter than the effective spin

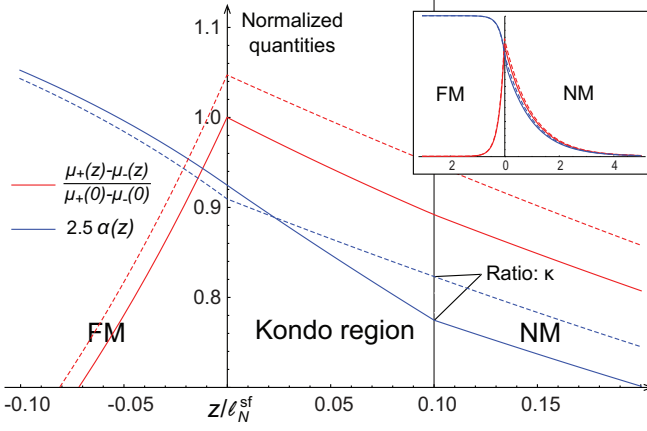


FIG. 2: (color online) Profiles of the spin accumulation $\mu_+(z) - \mu_-(z)$ divided by its value at $z = 0$ (red), and the current polarization $\alpha(z) = [j_+(z) - j_-(z)]/j_{\text{app}}$ multiplied by a factor of 2.5 (blue) to allow it to be plotted on the same scale as the spin accumulation. The dashed lines denote the solutions without dilute magnetic impurities with the same normalization factors. κ is defined as the ratio of $\alpha(d)$ for the case of finite magnetic impurity concentration to that when no magnetic impurities are present [Eq. (8)]. κ defined by the spin accumulation and that defined by the spin current have the same value. The inset shows the profiles over a wider range $z/l_N^{\text{sf}} = -3$ to 5. The parameters used here are $\sigma_{\pm,F} = (0.1 \pm 0.05)\sigma_N$, $l_N^{\text{sf}} = 5l_F^{\text{sf}} = 10d$, $\tilde{l}_N^{\text{sf}}/l_N^{\text{sf}} = 0.9$, and $\tilde{\sigma}_N/\sigma_N = 0.7$.

diffusion length. After straightforward but tedious algebra, we obtain

$$\kappa = 1 - d \left[\frac{1 + 2\eta(\tau_N^{\text{sf}}/\tau_N)}{l_N^{\text{sf}}} - \frac{2\eta(\tau_N^{\text{sf}}/\tau_N)}{l_N^{\text{sf}} + \frac{\rho_F}{(1-\alpha_{\text{FM}}^2)\rho_N} l_F^{\text{sf}}} \right] \frac{\rho_K}{\rho_N}, \quad (9)$$

where ρ_K is the Kondo contribution to the resistivity, $\rho_N = (2\sigma_N)^{-1}$ is the electrical resistivity of the normal metal, and $\rho_F = (\sigma_{+,F} + \sigma_{-,F})^{-1}$ is the electrical resistivity of the ferromagnet. From the Drude model, $\rho_K/\rho_N = \tau_N/\tau_K^{\text{eff}}$. $\alpha_{\text{FM}} = \alpha(z = -\infty)$, the current polarization far away from the interface, is a material parameter determined by the conductivity polarization $(\sigma_{+,F} - \sigma_{-,F})/(\sigma_{+,F} + \sigma_{-,F})$. The advantage of writing Eq. (9) in terms of ρ_K instead of τ_K^{eff} is that we can avoid the original Kondo expression $\sim \log T$, which diverges at low temperature, and instead use the phenomenological expression for ρ_K suggested by Goldhaber-Gordon [36] [Eq. (11)], which is known to work well for a wide range of temperatures [37, 38]. If $\tau_N^{\text{sf}} \gg \tau_N$, $1 + 2\eta(\tau_N^{\text{sf}}/\tau_N) \approx 2\eta(\tau_N^{\text{sf}}/\tau_N)$, thus one can verify that $1 - \kappa$ is proportional to $(\tau_N^{\text{sf}}/\tau_N)$, which is on the order of 10^3 for Cu. Such a large factor shows that the spin diffusion length is indeed a good tool to observe the Kondo effect, as discussed in Sec. II A.

For an order-of-magnitude estimate of the suppression ratio $1 - \kappa$ we take: $l_N^{\text{sf}} \approx 500$ nm and $l_F^{\text{sf}} \approx 10$ nm [1], with $\rho_F/(1 - \alpha_{\text{FM}}^2)\rho_N \approx 10$ and $2\eta\tau_N^{\text{sf}}/\tau_N \approx 10^3$. With $\rho_K/\rho_N \approx 0.01$, $1 - \kappa$ is around a percent for $d = 1$ nm and is proportional to d . This crude order-of-magnitude estimation is comparable to our previous experiments [29, 30].

III. DETAILED COMPARISON WITH EXPERIMENTAL DATA

The suppression of the spin polarization α [quantified by $1 - \kappa$ in Eq. (9)] has been experimentally observed in nonlocal spin valves fabricated from a variety of miscible, moment-forming ferromagnet/nonmagnet pairings, e.g., Fe/Cu and $\text{Ni}_{80}\text{Fe}_{20}/\text{Cu}$ [29]. Recently, through the use of thermal annealing to promote interdiffusion, we have also shown fine control over κ in Fe/Cu nonlocal spin valves, and directly correlated its magnitude to the Fe/Cu interdiffusion length, λ_{Fe} [30]. Equation (9) quantitatively connects $\kappa(T)$ to measurable quantities for the first time, and so provides an ideal expression with which to compare to these experiments. In the current section, we examine the experimental magnitude and T dependence of κ , while varying the annealing temperature, T_A , in order to tune the extent of the interfacial Kondo region. Through this analysis, and the use of Eq. (9), we extract an experimental value for the Elliot-Yafet probability for Kondo spin-flip scattering η , demonstrating good agreement between the presented theory and experimental results.

In nonlocal spin valve measurements, where ferromagnetic contacts are separated by a distance L , $\alpha(T)$ is typically extracted by fitting $\Delta R_{\text{NL}}(L, T)$ to a one-dimensional model of nonlocal spin transport [39]. In this model α_{eff} enters as a boundary condition which principally determines the magnitude of ΔR_{NL} at fixed L . Provided that the Kondo region is small compared with the mesoscopic device length ($d < L$), Kondo depolarization then appears as an interfacial effect, and is manifest as a suppression of the measured α at low T , as quantified by κ [Eq. (9)]. To account for this, we define an effective polarization, $\alpha_{\text{eff}} = \alpha(d)$, i.e., the observed current polarization in nonlocal spin valves with Kondo suppression present. This can be contrasted with the intrinsic polarization of the ferromagnet, α_{FM} . In this context, $\kappa \approx \alpha_{\text{eff}}/\alpha_{\text{FM}}$, and so determining $\alpha_{\text{eff}}(T)$, through $\Delta R_{\text{NL}}(L)$, and $\alpha_{\text{FM}}(T)$ yields an experimental measure of $\kappa(T)$. In the following, we examine $\alpha_{\text{eff}}(T)$ obtained from annealed Fe/Cu nonlocal spin valves. (Details of experimental fabrication and measurement of these all-metallic nonlocal spin valves can be found in the original reports). We note that Fe/Cu represents an ideal choice of materials as Fe is miscible and moment forming in Cu [40], with a readily accessible $T_K = 30$ K [41].

To determine $\alpha_{\text{FM}}(T)$ we also measured $\Delta R_{\text{NL}}(L, T)$ in devices devoid of dilute impurity moments, and thus the Kondo effect. Two types of devices were tested along these lines: nonlocal spin valves fabricated from nonmagnets that do not support local moments, e.g., Al; and nonlocal spin valves that incorporate a thin interlayer (e.g., Al) between the ferromagnet and nonmagnet that suppresses interdiffusion and moment formation. In both types of device the normalized $\alpha_{\text{FM}}(T)$ is found to be monotonic and quantitatively similar, as shown in black squares in Fig. 3(a).

To avoid complications from potential interface resistance changes during annealing, as well as other inherent systematic errors between devices, we scale α_{eff} to $\alpha_{\text{FM}}(T)$ using the method discussed in Appendix B. The resulting normalized α_{eff} for various T_A are shown in Fig. 3(a), with the corre-

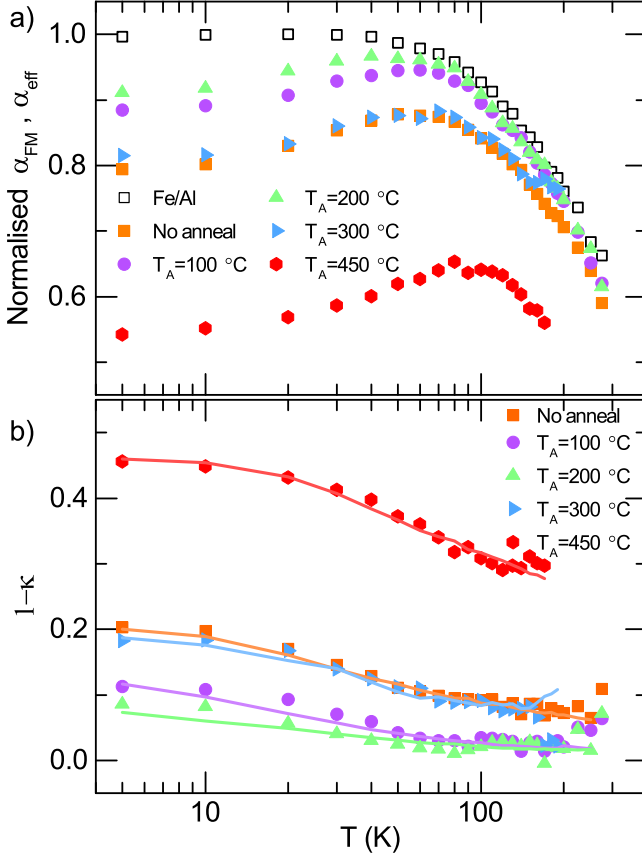


FIG. 3: (color online) (a) Temperature dependence of the normalized α_{FM} , obtained from Fe/Al nonlocal spin valves (open symbols), and α_{eff} for Fe/Cu nonlocal spin valves annealed at various temperatures (closed symbols), T_A . α_{eff} is normalized to α_{FM} using the procedure described in Appendix B. (b) Extracted $1 - \kappa(T)$ from α_{eff} . Symbols represent experimental data for various T_A . Solid lines in (b) are fits to the data using Eq. (9), with unconstrained T_K and a phenomenological Goldhaber-Gordon expression for ρ_K .

sponding $1 - \kappa(T)$, i.e. the degree of suppression, shown in Fig. 3(b). Each dataset here comes from fitting $\Delta R_{\text{NL}}(L, T)$ from devices with at least four different contact separations, ranging from $L = 250$ nm to 5 μm . The data of Fig. 3(a) are taken from a larger set of nonlocal spin valves (eight in total), however, for clarity we show only one curve at each T_A . For devices that do not support local moments $\alpha(T)$ is found to monotonically decrease with increasing T . In the presence of interdiffusion, however, a suppression of $\alpha(T)$ is observed at low T , the magnitude of which broadly increases with increasing T_A . Consequently, $1 - \kappa(T)$ is largest at high T_A and decreases with increasing measurement T , as would be anticipated. Figure 4(a) shows $1 - \kappa$ at $T = 5$ K for all samples, demonstrating this increase in magnitude with T_A . Note Fig. 4 displays data for all eight measured device batches, including multiple sets at $T_A = 200^\circ\text{C}$ and 300°C .

We now consider fitting the data of Fig. 3(b) using Eq. (9). In this equation $\rho_N^{\text{sf}}(T)$, $\rho_N(T)$ and $\rho_F(T)$ are measured experimentally. $\tau_N^{\text{sf}}/\tau_N$ is constrained to the literature value

of 950 [42] — a value we have explicitly verified in high-conductivity nonlocal spin valve channels, where Kondo effects are negligible and phonon scattering dominates spin relaxation. We consider two expressions for $\rho_K(T)$, the Kondo model [31] [Eq. (10)] and the phenomenological formalism of Goldhaber-Gordon (G-G) [36] [Eq. (11)]:

$$\rho_K^K = \rho_m \left(1 + 2N_0 J \ln \frac{T}{T^*} \right), \quad (10)$$

$$\rho_K^{\text{G-G}} = \rho_m \left[1 + 2N_0 J \left(\frac{T_K'^2}{T^2 + T_K'^2} \right)^s \ln \frac{U}{k_B T^*} \right], \quad (11)$$

where $T_K' = T_K / \sqrt{2^{1/s} - 1}$, $\rho_m = \hat{C}_{\text{Fe}} 2\pi\mu_N^0 S(S+1)N_0^2 J^2 m / 3\hbar n_N^0 e^2$ is the classical resistivity without taking account of dilute magnetic impurities, U is the on-site Coulomb energy, N_0 is the density of states of each spin band, J is the (negative) exchange parameter between the conduction electron and the magnetic impurities, and s is the so-called G-G exponent which is typically taken to be 0.22 [36]. In the expression for ρ_m , μ_N^0 is the Fermi level of the normal metal, \hat{C}_{Fe} is the average impurity concentration in Fe, S is the spin angular momentum of the magnetic impurities, n_N^0 is the density of electrons, m is the effective electron mass. Here we have adapted the generalized G-G model to give agreement with Kondo's original theory. T^* depends on the limits of the energy integration of Eq. (24) and is typically taken to be either T_K , U or J , depending on the theoretical treatment. In our case, the requirement that both models be equivalent at $T = T_K$ gives $T^* = k_B T_K^2 / U$. The Fermi energy of Cu, $\mu_N^0 = 7$ eV [43], is well known, as is $T_K = 30$ K from $\rho_K(T)$ and susceptibility measurements [41]. Furthermore, $J = 0.91$ eV for Fe/Cu has been experimentally measured via field-dependent magneto-resistance and magnetometry measurements [44]. From our own measurements of $\rho_K(T)$ in heavily doped Fe/Cu nanowires, we can establish $U \approx 0.86$ meV (10 K). Noting $\tau_N^{\text{sf}}/\tau_N \gg 1$, this leaves only the product $\eta \hat{C}_{\text{Fe}} d$, i.e., the weighted total number of impurities in the Kondo region (per cross sectional area), as an unknown. Equation (2) is known to be valid only over a limited T range about T_K , evolving to the value dictated by the classical scattering rate at $T \gg T_K$ and the unitary limit at $T \ll T_K$. Consequently, when using this model we restrict the fit T range to only consider the transition region within the data, between approximately 10 K and 100 K. In addition to considering both of these models, we also compare to the cases where T_K is allowed to be an unconstrained fitting parameter. The extracted $\eta \hat{C}_{\text{Fe}} d$ from each model (through a least mean square minimization of residuals) is shown in Fig 4(b). We note that the individual parameters η , \hat{C}_{Fe} and d remain otherwise inseparable. The solid lines in Figure 3(b) show fits of $1 - \kappa(T)$ using the G-G model with unconstrained T_K . In general, the overall magnitude and T dependence is well captured for all T_A . When T_K is unconstrained we find $T_K = (44 \pm 36)$ K for the G-G model and $T_K = (22 \pm 9)$ K for the Kondo model, in good agreement with the literature value of $T_K = 30$ K. (All uncertainties in this paper are single standard deviations, the determination of which is discussed in Appendix D.) The deviation in these values reflects the logarithmic dependence

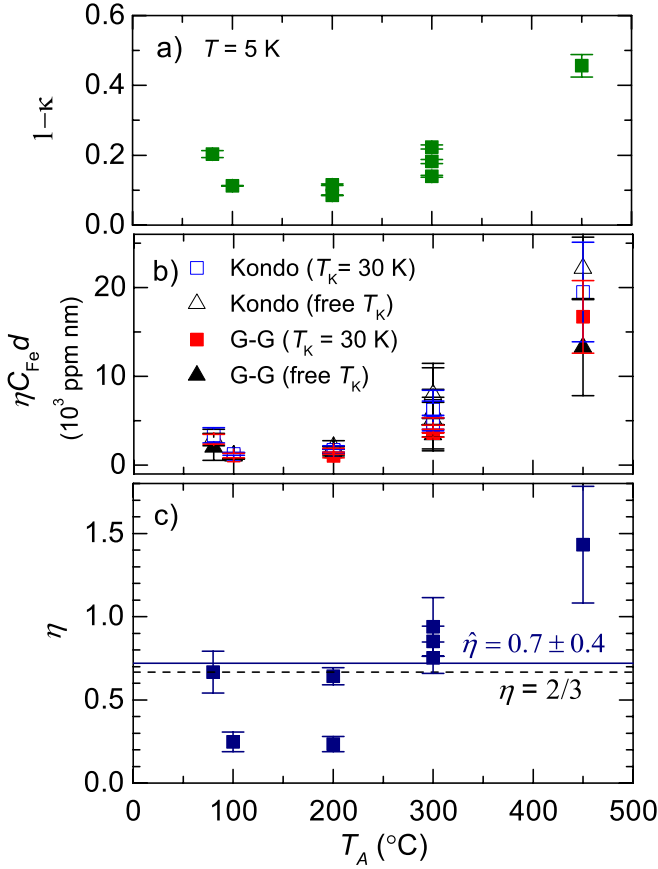


FIG. 4: (color online) (a) Magnitude of $1 - \kappa$ for low T ($=5$ K) as a function of annealing temperature, T_A . (b) Extracted values of $\eta \hat{C}_{Fe} d$ using either the Kondo (open symbols) or Goldhaber-Gordon (G-G) expressions (closed symbols) for ρ_K , with T_K constrained (colored) or unconstrained (gray). (c) Estimated values of η using the results of panel b, with $\eta \hat{C}_{Fe} d = 560 \mu\text{mol/mol} \cdot \lambda_{Fe}$ and λ_{Fe} determined from prior STEM/EDX measurements. The error bars indicate single standard deviation uncertainties as discussed in Appendix D.

of ρ_K on T , and so the difficulty in determining T_K , which is often a challenge in dilute-moment metallic systems. Despite the limited range of applicability, both expressions (i.e., for ρ_K^K and ρ_K^{G-G}) give consistent results for $\eta \hat{C}_{Fe} d$, reflecting the fact that these parameters only act to influence the *magnitude* of κ , and are thus relatively insensitive to the precise T dependence of the data. Examining Fig. 4(b) we see that the total number of impurities increases on annealing, with dramatic changes occurring above $T_A \approx 300$ °C, in good agreement with our previous observations of the T_A dependence of λ_{Fe} in this system [30].

To place these values of $\eta \hat{C}_{Fe} d$ in context, and to extract a value of η , we consider our previous work on interdiffusion in Fe/Cu nonlocal spin valves. Fe has finite solubility in Cu, with a limit of $2600 \mu\text{mol/mol}$ at room temperature [40] (based on the bulk equilibrium phase diagram), beyond which precipitation occurs, leading to phase-segregated clusters. In the following analysis we therefore assume that regions with $\hat{C}_{Fe} > 2600 \mu\text{mol/mol}$ do not contain isolated

dilute moments, and so do not contribute to the Kondo effect. Through energy-dispersive x-ray spectroscopy (EDX) measurements in cross-sectional scanning transmission electron microscopy (STEM) we have previously shown that the interdiffusion profile in annealed Fe/Cu nonlocal spin valves follows $C_{Fe}(x) \propto [1 - \text{erf}(x/\lambda_{Fe})]/2$, and have quantitatively determined $\lambda_{Fe}(T_A)$ for our devices [30]. Using this expression for C_{Fe} , considering only the dilute Kondo region below $C_{Fe} < 2600 \mu\text{mol/mol}$ in nonlocal spin valves, and assuming $L \gg \lambda_{Fe}$, yields a total number of impurity atoms [50] (per unit cross sectional area) in the Kondo region of $\hat{C}_{Fe} d = 560 \mu\text{mol/mol} \cdot \lambda_{Fe}$. Using this result and our previous measurements of λ_{Fe} , we are therefore able to estimate η . Duly extracted values of η are shown in Fig. 4(c) as a function of T_A , with the average $\hat{\eta} = 0.7 \pm 0.4$ indicated by the solid horizontal line. We anticipate that the simplicity of the model in accounting for the precise dispersion of Fe at the Fe/Cu interface, in addition to variability in the precise degree of interdiffusion at intermediate T_A , likely accounts for the dispersion in the extracted values of η , particularly at $T_A = 200$ °C. This is also likely to be the cause of the unphysical value of $\eta > 1$ found at $T_A = 450$ °C (note also the large random error). Despite the variation, we find good overall consistency between the experimentally determined values and $\eta = 2/3$ (dashed horizontal line), as calculated in the original work of Kondo [31]. Given the rather simplistic assumptions made, as well as the use of four independent experimental measurements in order to determine $\hat{\eta}$, this result indicates good consistency between the model and experiment.

As a final point of discussion, we note that one of the most striking observations of Ref. [30] was the failure of the standard one-dimensional Valet-Fert model to describe the temperature and separation dependence of ΔR_{NL} at high C_{Fe} . It is important to highlight that, due to heavy interdiffusion, this situation involves not only a large C_{Fe} , but also a large gradient over relatively long distances. The conclusion from this should not be that the Valet-Fert theory itself is fundamentally incapable of describing the impact of Kondo relaxation on spin transport. Rather, it is the application of models, based on Valet-Fert theory but which, crucially, ignore Kondo physics, which fails to describe even qualitative observations. The results in this work suggest, in fact, that it is indeed possible for an (albeit modified) Valet-Fert approach to describe previous experimental results, when the precise distribution of local moments throughout the device is known.

IV. DERIVATIONS

In this section we present mathematical derivations of the core results in Sec. II A. First we start from the Boltzmann equation to derive the spin drift-diffusion equation at finite temperature. We take the approach developed by Valet and Fert [17]. Readers not familiar with details of the Valet-Fert theory may refer to Appendix C 1.

We start from the Boltzmann equation for a translation-invariant system in the two directions perpendicular to z . The distribution functions f_{\pm} are a function of z and \mathbf{k} . At the

equilibrium, they are $[f_s^0(z, k)]^{-1} = e^{\beta[\varepsilon(\mathbf{k}) - \mu_s^0(z)]} + 1$ where $\varepsilon(\mathbf{k}) = \hbar^2 k^2 / 2m$ is the energy eigenvalue, \mathbf{k} is the crystal momentum, $k = |\mathbf{k}|$, $\mu_s^0(z)$ is the chemical potential at the equilibrium, m is the effective mass, $\beta = 1/k_B T$ is the inverse temperature, and k_B is the Boltzmann constant. When an electric field is applied along the \hat{z} direction, the distribution function has a small correction to its equilibrium function $f_s(z, \mathbf{k}) = f_s^0(\mathbf{k}) + g_s(z, \mathbf{k})$ where $g_s(\mathbf{k})$ is the small correction proportional to the electric field and $s = \pm$. The linearized Boltzmann equation at steady state under the relaxation time approximation is

$$\begin{aligned} & \frac{\hbar k_z}{m} \partial_z g_s(z, \mathbf{k}) - \frac{eE}{\hbar} \partial_{k_z} f_s^0(k) \\ &= -\frac{g_s(z, \mathbf{k}) - g_s^{(0)}(z, k)}{\tau_s^{\text{sc}}(z, k)} - \frac{g_s(z, \mathbf{k}) - g_{-s}^{(0)}(z, k)}{\tau_s^{\text{sf}}(z, k)}, \end{aligned} \quad (12)$$

where τ_s^{sc} and τ_s^{sf} are the relaxation times. τ_s^{sc} corresponds to spin-conserving processes $s \rightarrow s$, and τ_s^{sf} corresponds to spin-flipping processes that gives the spin diffusion length. A two-channel model works well in materials with weak spin-orbit coupling, like Cu, where the different cartesian components of the spin are largely decoupled from one another. This allows us to neglect spin components perpendicular to the magnetization in the ferromagnet in our derivation. The sum of the spin-conserving and the spin-flipping rates gives the electrical conductivity described below. These relaxation times are assumed to be isotropic, that is, they only depend on the magnitude of k . In general τ_s^{sf} is independent of s at the Fermi surface. However, since the Fermi wave vector depends on s , τ_s^{sf} written as a function of k is s dependent. This is also true for the Kondo contribution [31]. The relaxation times τ_N and τ_N^{sf} that appear in Sec. II will be connected to these functions. $g_s^{(0)}(z, k)$ is the angle-averaged $g_s(z, \mathbf{k})$ over the Fermi surface with the constant magnitude $|\mathbf{k}| = k$, that is, $g_s^{(0)}(z, k) = (1/4\pi) \int_{\Omega_{\mathbf{k}}} d\Omega_{\mathbf{k}} g_s(z, \mathbf{k})$ where $\Omega_{\mathbf{k}}$ is the solid angle of \mathbf{k} .

We make the approximation that the system is rotationally symmetric around the z axis. In this regime, $g(z, \mathbf{k})$ can be expanded by the Legendre polynomials as $g_s(z, \mathbf{k}) = \sum_{n=0}^{\infty} g_s^{(n)}(z, k) P_n(\hat{\mathbf{k}} \cdot \hat{\mathbf{z}})$, where P_n is the Legendre polynomial. Since the Legendre polynomials form an orthonormal set of polynomials, each coefficient satisfies the equation. Taking $n = 0, 1$ coefficients and neglecting the higher order contributions [17], Eq. (12) is equivalent to

$$\frac{\hbar k}{3m} \partial_z g_s^{(1)}(z, k) = -\frac{\tilde{g}_s^{(0)}(z, k) - \tilde{g}_{-s}^{(0)}(z, k)}{\tau_s^{\text{sf}}(z, k)}, \quad (13)$$

$$\frac{\hbar k}{m} \partial_z \tilde{g}_s^{(0)}(z, k) = -\frac{g_s^{(1)}(z, k)}{\tau_s(z, k)}, \quad (14)$$

where $\tilde{g}_s^{(0)} = g_s^{(0)} - eEz\partial_{\varepsilon} f_s^0$, $1/\tau_s = 1/\tau_s^{\text{sc}} + 1/\tau_s^{\text{sf}}$ is the total scattering-out rate of a spin s state.

Equations (13) and (14) are the spin drift-diffusion equations that hold at $T = 0$. At $T = 0$, $\tilde{g}_s^{(0)}$ and $g_s^{(1)}$ evaluated at the Fermi surface are respectively assigned to the chemical potential and the current, with proper prefactors [17]. Therefore,

Eqs. (13) and (14) provide closed solutions for these physical quantities. However, this association is not exact at finite temperature. For $T > 0$, a physical quantity is not given by a value at the Fermi surface, but is given after integrating over k , considering T dependence of $f_s^0(k)$.

Although the temperature dependence of f_s^0 is very complicated, the Sommerfeld expansion formula in Appendix C 2 allows substantial simplification. The Sommerfeld expansion formula is an expression for integrals including the Fermi-Dirac distribution at low temperature. In the Kondo regime, $k_B T \ll \mu_s^0$ satisfies the criterion. Neglecting $O(T^2)$, the spin density n_s for each s band (or the spin chemical potential with an additional factor as shown below) is given by

$$\begin{aligned} n_s(z) &= \frac{e}{V} \sum_{\mathbf{k}} f_s(z, \mathbf{k}) \\ &= \frac{e}{2\pi^2} \int dk k^2 \tilde{g}_s^{(0)}(z, k) \\ &\quad + \frac{m \sqrt{m \mu_s^0(z)} e^2 E z}{\sqrt{2} \pi^2 \hbar^3} + \frac{m \sqrt{2m \mu_s^0(z)} \mu_s^0(z) e}{3\pi^2 \hbar^3}. \end{aligned} \quad (15)$$

We use here the Sommerfeld expansion formula Eq. (C7) for low temperature $k_B T \ll \mu_s^0$. Similarly, the current density $j_s(z)$ for s band is

$$j_s(z) = -\frac{e}{V} \sum_{\mathbf{k}} \frac{\hbar k_z}{m} f_s(z, \mathbf{k}) = -\frac{e\hbar}{6\pi^2 m} \int dk k^3 g_s^{(1)}(z, k). \quad (16)$$

Assuming that $\mu_s^0(z)$ is constant in space in each region (sudden changes at the boundaries can be taken into account by matching boundary conditions), integration of Eqs. (13) and (14) with the weighting factors k^2 and k gives respectively

$$\partial_z j_s(z) = \frac{e}{2\pi^2} \int dk k^2 \frac{\tilde{g}_s^{(0)}(z, k) - \tilde{g}_{-s}^{(0)}(z, k)}{\tau_s^{\text{sf}}(z, k)}, \quad (17)$$

$$\partial_z n_s(z) = -\frac{em}{2\pi^2 \hbar} \int dk k \frac{g_s^{(1)}(z, k)}{\tau_s(z, k)}. \quad (18)$$

Equations (17) and (18) provide a generalized drift-diffusion equation at low temperature.

We first briefly show that Eqs. (17) and (18) become the conventional spin drift-diffusion equations without the Kondo effect. That is, Valet-Fert theory holds up to $O(T)$, if there are no magnetic impurities. Without an electric field, $\tilde{g}_s^{(0)}$ and $g_s^{(1)}$ are zero. This observation implies that $\tilde{g}_s^{(0)}$ and $g_s^{(1)}$ are proportional to $\partial_{\varepsilon} f_s^0$. By the Sommerfeld expansion formula, $\partial_{\varepsilon} f_s^0$ can be replaced by $-\delta(\mu_s^0 - \varepsilon)$, neglecting $O(T^2)$. Thus, the integrations Eqs. (17) and (18) are nothing but evaluations at the Fermi surface. Therefore, Eqs. (17) and (18) are equivalent to Eqs. (13) and (14) up to $O(T)$.

Now we connect the quantities that appear in Eqs. (13) and (14) to physical quantities. First we define the electrical conductivity $\sigma_s(z)$ and the spin diffusion length $l_s(z)$ for each spin band s , which are respectively given by

$$\sigma_s(z) = \frac{e^2 (k_s^F)^3 \tau_s(z, k_s^F)}{6\pi^2 m}, \quad (19)$$

$$l_s^2(z) = \frac{\hbar^2 (k_s^F)^2}{3m^2} \tau_s(z, k_s^F) \tau_s^{\text{sf}}(z, k_s^F), \quad (20)$$

where $k_s^F = \sqrt{2m\mu_s^0/\hbar^2}$ is the Fermi wave vector. Here the electrical conductivity is equivalent to the Drude conductivity $\sigma_s = n_s^0 e^2 \tau_s(k_s^F)/m$ where $n_s^0 = (k_s^F)^3/6\pi^2$ is n_s without an electric field. Similarly, the spin diffusion length is related to the diffusion constant by $l_s^2 = D_s \tau_s^{\text{sf}}(k_s^F)$ where $D_s = \hbar^2 (k_s^F)^2 \tau_s(z, k_s^F)/3m^2$. Next, we define the electrochemical potential $\mu_s = (2\pi^2 \hbar^2 / em k_s^F) n_s$. The factor arises from the ratio between $\int d\varepsilon$ and $(e/V) \sum_{\mathbf{k}}$ in Eq. (15). With these definitions, Eqs. (13) and (14) become equivalent to Eqs. (5) and (6).

The situation changes in the presence of dilute magnetic impurities. We show that Eqs. (5) and (6) are still valid after replacement of Eqs. (3) and (4) [or equivalently Eqs. (1) and (2)] and give explicit expressions for τ_K^{eff} in particular regimes. Since the Kondo effect occurs in the normal metal, we use the subscript N but drop the spin-dependent subscript s . That is, $\tau_N(z, k)$ and $\tau_N^{\text{sf}}(z, k)$ are the relaxation times in the normal metal, which are spin independent. τ_N and τ_N^{sf} that appear in Sec. II are those evaluated at the Fermi level $k = k_N^F$. In the presence of dilute magnetic impurities, additional relaxation rates arise due to the impurities. We denote these by the subscript K . In the Boltzmann equation Eq. (12), the relaxation times change by

$$\frac{1}{\tilde{\tau}_N^{\text{sc}}(z, k)} = \frac{1}{\tau_N^{\text{sc}}(z, k)} + \frac{1}{\tau_K^{\text{sc}}(z, k)}, \quad (21)$$

$$\frac{1}{\tilde{\tau}_N^{\text{sf}}(z, k)} = \frac{1}{\tau_N^{\text{sf}}(z, k)} + \frac{1}{\tau_K^{\text{sf}}(z, k)}, \quad (22)$$

where $\tilde{\tau}_N^{\text{sc}}$ and $\tilde{\tau}_N^{\text{sf}}$ are the modified relaxation times due to the Kondo effect. Kondo [31] computed explicitly the total scattering rate change $\tau_K^{-1} = (\tau_K^{\text{sc}})^{-1} + (\tau_K^{\text{sf}})^{-1}$ given by

$$\frac{1}{\tau_K(z, k)} = \frac{2\pi\mu_N^0 S(S+1)\hat{C}N_0^2 J^2}{3\hbar} [1 + 2J\gamma(\varepsilon)], \quad (23)$$

$$\gamma(\varepsilon) = \frac{1}{V} \sum_{\mathbf{k}} \frac{f_N^0(\mathbf{k})}{\varepsilon(\mathbf{k}) - \varepsilon}, \quad (24)$$

where S is the spin angular momentum of the magnetic impurities, $N_0 = mk_N^F/2\pi^2\hbar^2$ is the density of states of each spin band, J is the (negative) exchange parameter between the conduction electron and the magnetic impurities, and \hat{C} is the average impurity concentration, which is the density of impurities divided by the density of electrons $2n_N^0 = (k_N^F)^3/3\pi^2$. The units of J are $\text{J} \cdot \text{m}^3$ to make $N_0 J$ dimensionless. To be explicit, we set the Kondo Hamiltonian to be $H_K = -(4J/V) \sum_{\mathbf{k}, \mathbf{d}} (\Psi_{\mathbf{k}}^\dagger \sigma \Psi_{\mathbf{k}}) \cdot (\Psi_{\mathbf{d}}^\dagger \sigma \Psi_{\mathbf{d}})$, where $\Psi_{\mathbf{k}}^\dagger$ and $\Psi_{\mathbf{d}}^\dagger$ are respectively the electron creation operator of conduction electrons with momentum \mathbf{k} and electrons in the impurity state \mathbf{d} and σ is the Pauli matrix.

Since the Kondo theory is a perturbation theory, its contribution to the spin-flip rate $(\tau_K^{\text{sf}})^{-1}$ is likely to be proportional to Eq. (23) as Kondo showed [31]. We introduce η , the spin-flip probability during each Kondo scattering event by

$$\frac{1}{\tau_K^{\text{sf}}(z, k)} = \frac{\eta}{\tau_K(z, k)}. \quad (25)$$

The value of η is determined by the Fermi surface geometry. For a spherical Fermi surface that we use here, Kondo [31] showed that $(\tau_K^{\text{sf}})^{-1}$ is twice $(\tau_K^{\text{sc}})^{-1}$ [51]. This indicates that $\eta = 2/3$ for this case.

The low temperature behavior of $\tau_K(z, k)$ requires careful treatment. Since $\gamma(\varepsilon)$ diverges at the Fermi level, naive application of the Sommerfeld expansion gives divergences. Although the integrals in Eqs. (17) and (18) seem surprisingly difficult to perform without the Sommerfeld expansion, a low temperature approximation allows it. In Appendix C3, we slightly generalize Kondo's approach to extract the logarithmic dependence of the Kondo resistivity to show that the following replacement is valid under the integration over the energy.

$$\gamma(\varepsilon) \partial_\varepsilon f_N^0 \rightarrow N_0 \ln \frac{k_B T}{\mu_N^0} [-\delta(\varepsilon - \mu_N^0)], \quad (26)$$

giving rise to a $\ln T$ contribution. With this rule, Eqs. (5) and (6) still hold under the following replacement.

$$\frac{1}{\tau_N(z, k_s^F)} \rightarrow \frac{1}{\tilde{\tau}_N(z, k_s^F)} \equiv \frac{1}{\tau_N(z, k_s^F)} + \frac{1}{\tau_K^{\text{eff}}}, \quad (27)$$

$$\frac{1}{\tau_N^{\text{sf}}(z, k_s^F)} \rightarrow \frac{1}{\tilde{\tau}_N^{\text{sf}}(z, k_s^F)} \equiv \frac{1}{\tau_N^{\text{sf}}(z, k_s^F)} + \frac{\eta}{\tau_K^{\text{eff}}}, \quad (28)$$

which are nothing but Eqs. (1) and (2). Here

$$\frac{1}{\tau_K^{\text{eff}}} = \frac{2\pi\mu_N^0 S(S+1)\hat{C}N_0^2 J^2}{3\hbar} \left(1 + 2N_0 J \ln \frac{k_B T}{\mu_N^0} \right). \quad (29)$$

V. SUMMARY

In order to take account for the Kondo effect in spin transport, we derive a modified drift-diffusion equation from the Boltzmann equation explicitly allowing for finite temperature. The complicated finite temperature theory is projected to a low temperature regime (compared to the Fermi temperature). We show that the Valet-Fert drift-diffusion equation holds both at finite T and in the presence of spin scattering from dilute magnetic impurities, once the electrical conductivity and spin diffusion length are renormalized as functions of temperature. This represents a useful result; as a consequence, dilute magnetic impurity scattering beyond the semiclassical limit can indeed be described in a simple Elliot-Yafet-like form with a direct proportionality between τ_K^{sf} and τ_K , as originally indicated by Kondo. The modified drift-diffusion equation has a remarkably compact form given the complexity of the higher order many-body interactions involved.

By solving the drift-diffusion equation for an illustrative regime, we show additional spin relaxation in the presence of the Kondo effect at a ferromagnet/nonmagnet interface. Kondo scattering is found to be highly efficient at spin relaxation, due to the high probability of spin flip ($\eta = 2/3$) compared with other scattering mechanisms (c.f. $\eta_{\text{phonon}} \approx 1/1000$). Since the spin-flip rate is much lower than the momentum scattering rate in the absence of the Kondo effect,

such a high probability caused by the Kondo effect can significantly reduce the spin diffusion length, even when there is negligible change to the conductivity. This is confirmed experimentally by the large value of $\eta \approx 0.7$ observed, in good agreement with Kondo's original work. We hope this, in addition to the explicit derivation of Eq. (4), further validates the semiclassical model of Ref. [28] in also determining the Kondo contributions to I_N . Note again that the fitting procedure used here relies on four independent quantities that are experimentally extracted, as well as approximations regarding the precise distribution of magnetic moments within the Kondo region. Examining Fig. 4, one can see a weak dependence of η on T_A . This is very likely due to such simplifications. Indeed, the possibilities of Fe segregation and cluster formation on annealing, dilute impurity migration to grain boundaries, and inter-moment correlations at high concentrations, as well as examining the precise phase equilibrium beyond the thermodynamic limit are entirely overlooked, and could greatly complicate the situation. Nevertheless, agreement between the simple model and experiment is highly satisfactory.

One observation worth mentioning is the form of Eq. (9), particularly the fact the signal suppression is linear in both \hat{C}_{Fe} and d . This clarifies one of the fundamental difficulties previously experienced within the field. That is, determining the precise location of the anomalous relaxation mechanism. Previous reports have stated relaxation occurring at the ferromagnet/nonmagnet interface (as we discuss here), throughout the channel [27, 28], or at surfaces [21, 26, 45], with similar magnitudes of Kondo suppression in each case. To first order it is the product $\hat{C}_{Fe}d$ (total number of impurities per cross-sectional area) that determines suppression, and so similar magnitudes may be observed *either* due to a high-impurity-concentration narrow region (e.g. an interfacial effect), or an extended low concentration region (i.e., low doping levels throughout the channel itself). For the case where magnetic impurities extend throughout the channel, i.e. in the limit where $d \geq \tilde{l}_N$, the approximations made in obtaining Eq. (9) will no longer be appropriate. Instead separation-dependent measurements of ΔR_{NL} on mesoscopic lengthscales (i.e. comparable to \tilde{l}_N) will follow the standard non-local spin trans-

port equations, now with the modified value of I_N given by Eq. 4. For low impurity levels this results in comparable magnitudes of suppression to those seen here. Rather than serendipitous, it is entirely expected that both the interfacial effects discussed here and 'contaminated' channel devices observe similar signal contributions from Kondo effects. This highlights the care that must be taken when fitting $\Delta R_{NL}(L, T)$ to resolve the contributions from interfacial (manifest in the extracted α) and bulk (manifest in I_N) Kondo relaxation, in the likely scenario where both cause ΔR_{NL} to be suppressed by comparable amounts.

Having now determined the theoretical and experimental T -dependence of Kondo spin scattering in nonlocal spin valves, this opens the path to using the Kondo effect to better understand magnetic and nonmagnetic impurity spin relaxation. In particular, the clearly identifiable T -dependence may now be used as a signature to quantitatively determine the contribution of dilute moments to relaxation in all-metal systems.

Acknowledgments

The authors acknowledge P. Haney and O. Gomonay for critical reading of the manuscript. K.W.K. was supported by the Cooperative Research Agreement between the University of Maryland and the National Institute of Standards and Technology, Center for Nanoscale Science and Technology (70NANB10H193), through the University of Maryland. K.W.K. also acknowledges support by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2016R1A6A3A03008831), Alexander von Humboldt Foundation, the ERC Synergy Grant SC2 (No. 610115), and the Transregional Collaborative Research Center (SFB/TRR) 173 SPIN+X. Work at the University of Minnesota was funded by Seagate Technology Inc., the University of Minnesota (UMN) NSF MRSEC under award DMR-1420013, and DMR-1507048. L.O'B. acknowledges a Marie Curie International Outgoing Fellowship within the 7th European Community Framework Programme (project No. 299376).

Appendix A: Solution of the spin drift-diffusion equation

In Ref. [17], the general solutions for Eqs. (5) and (6) are given by

$$\mu_+(z) - \mu_-(z) = \begin{cases} A_F e^{z/l_F^{\text{sf}}} & \text{for } z < 0, \\ \tilde{A}_N e^{z/\tilde{l}_N^{\text{sf}}} + \tilde{B}_N e^{-z/\tilde{l}_N^{\text{sf}}} & \text{for } 0 < z < d, \\ B_N e^{-z/l_N^{\text{sf}}} & \text{for } z > d, \end{cases} \quad (\text{A1})$$

$$\sigma_+(z)\mu_+(z) + \sigma_-(z)\mu_-(z) = \begin{cases} C_F z + D_F & \text{for } z < 0, \\ \tilde{C}_N z + \tilde{D}_N & \text{for } 0 < z < d, \\ C_N z + D_N & \text{for } z > d. \end{cases} \quad (\text{A2})$$

In this section we determine the coefficients satisfying the transparent boundary conditions

$$\mu_s(z = -0) = \mu_s(z = +0), \mu_s(z = d - 0) = \mu_s(z = d + 0), j_s(z = -0) = j_s(z = +0), j_s(z = d - 0) = j_s(z = d + 0). \quad (\text{A3})$$

There are eight boundary conditions (note that $s = \pm$) although there are ten unknown coefficients. Therefore, two more conditions are required. The first one originates from a constant shift of the chemical potential. Since the drift-diffusion equation is invariant under a constant shift of the chemical potential, we can put $\tilde{D}_N = 0$ without any loss of generality. The second one originates from the homogeneity of the drift-diffusion equation. The drift-diffusion equation is invariant under multiplication by a constant factor to μ_s . The applied electrical current defined by

$$ej_{\text{app}} = ej_+(z = -\infty) + ej_-(z = -\infty) = C_F, \quad (\text{A4})$$

is an experimentally controllable quantity that fixes the multiplication factor.

Now we apply the boundary conditions. Instead of applying the continuity of each functions, we can apply it with their independent linear combinations. Note that $\mu_+(z) - \mu_-(z)$ is already given above and $j_+(z) + j_-(z)$ is nothing but the derivative of $\sigma_+(z)\mu_+(z) + \sigma_-(z)\mu_-(z)$. Continuity of these functions at $z = 0$ and $z = d$ gives the following four conditions.

$$A_F = \tilde{A}_N + \tilde{B}_N, \quad \tilde{A}_N e^{d/\tilde{l}_N^{\text{sf}}} + \tilde{B}_N e^{-d/\tilde{l}_N^{\text{sf}}} = B_N e^{-d/\tilde{l}_N^{\text{sf}}}, \quad \tilde{C}_N = C_N = ej_{\text{app}}. \quad (\text{A5})$$

We now put these into the solution and obtain

$$\mu_+(z) - \mu_-(z) = \begin{cases} (\tilde{A}_N + \tilde{B}_N)e^{z/\tilde{l}_F^{\text{sf}}} & \text{for } z < 0, \\ \tilde{A}_N e^{z/\tilde{l}_N^{\text{sf}}} + \tilde{B}_N e^{-z/\tilde{l}_N^{\text{sf}}} & \text{for } 0 < z < d, \\ (\tilde{A}_N e^{d/\tilde{l}_N^{\text{sf}}} + \tilde{B}_N e^{-d/\tilde{l}_N^{\text{sf}}})e^{-(z-d)/\tilde{l}_N^{\text{sf}}} & \text{for } z > d, \end{cases} \quad (\text{A6})$$

$$\sigma_+(z)\mu_+(z) + \sigma_-(z)\mu_-(z) = \begin{cases} ej_{\text{app}}z + D_F & \text{for } z < 0, \\ ej_{\text{app}}z & \text{for } 0 < z < d, \\ ej_{\text{app}}z + D_N & \text{for } z > d. \end{cases} \quad (\text{A7})$$

Now we apply the continuity of $\mu_+(z) + \mu_-(z)$. After some algebra,

$$\mu_+(z) + \mu_-(z) = \begin{cases} \frac{2ej_{\text{app}}z + 2D_F}{\sigma_{+,F} + \sigma_{-,F}} - \frac{\sigma_{+,F} - \sigma_{-,F}}{\sigma_{+,F} + \sigma_{-,F}}(\tilde{A}_N + \tilde{B}_N)e^{z/\tilde{l}_F^{\text{sf}}} & \text{for } z < 0, \\ \frac{ej_{\text{app}}z}{\tilde{\sigma}_N} & \text{for } 0 < z < d, \\ \frac{ej_{\text{app}}z + D_N}{\sigma_N} & \text{for } z > d. \end{cases} \quad (\text{A8})$$

Continuity at $z = 0$ and $z = d$ gives

$$D_F = \frac{\sigma_{+,F} - \sigma_{-,F}}{2}(\tilde{A}_N + \tilde{B}_N), \quad D_N = \left(\frac{\sigma_N}{\tilde{\sigma}_N} - 1\right)ej_{\text{app}}d. \quad (\text{A9})$$

Then \tilde{A}_N and \tilde{B}_N are the only remaining coefficients. We now apply continuity of $j_+(z) - j_-(z)$. After some algebra,

$$\sigma_+(z)\mu_-(z) - \sigma_-(z)\mu_+(z) = \begin{cases} \frac{\sigma_{+,F} - \sigma_{-,F}}{\sigma_{+,F} + \sigma_{-,F}}(ej_{\text{app}}z + D_F) + \frac{2\sigma_{+,F}\sigma_{-,F}}{\sigma_{+,F} + \sigma_{-,F}}(\tilde{A}_N + \tilde{B}_N)e^{z/\tilde{l}_F^{\text{sf}}} & \text{for } z < 0, \\ \tilde{\sigma}_N(\tilde{A}_N e^{z/\tilde{l}_N^{\text{sf}}} + \tilde{B}_N e^{-z/\tilde{l}_N^{\text{sf}}}) & \text{for } 0 < z < d, \\ \sigma_N(\tilde{A}_N e^{d/\tilde{l}_N^{\text{sf}}} + \tilde{B}_N e^{-d/\tilde{l}_N^{\text{sf}}})e^{-(z-d)/\tilde{l}_N^{\text{sf}}} & \text{for } z > d. \end{cases} \quad (\text{A10})$$

Continuity of the derivatives at $z = 0$ and $z = d$ gives

$$\frac{\sigma_{+,F} - \sigma_{-,F}}{\sigma_{+,F} + \sigma_{-,F}}ej_{\text{app}} + \frac{2\sigma_{+,F}\sigma_{-,F}}{\sigma_{+,F} + \sigma_{-,F}}\frac{\tilde{A}_N + \tilde{B}_N}{\tilde{l}_F^{\text{sf}}} = \frac{\tilde{\sigma}_N}{\tilde{l}_N^{\text{sf}}}(\tilde{A}_N - \tilde{B}_N), \quad (\text{A11})$$

$$\frac{\tilde{\sigma}_N}{\tilde{l}_N^{\text{sf}}}(\tilde{A}_N e^{d/\tilde{l}_N^{\text{sf}}} - \tilde{B}_N e^{-d/\tilde{l}_N^{\text{sf}}}) = -\frac{\sigma_N}{\tilde{l}_N^{\text{sf}}}(\tilde{A}_N e^{d/\tilde{l}_N^{\text{sf}}} + \tilde{B}_N e^{-d/\tilde{l}_N^{\text{sf}}}), \quad (\text{A12})$$

the solutions of which are

$$\tilde{A}_N = e^{-2d/\tilde{l}_N^{\text{sf}}} \frac{\tilde{l}_N^{\text{sf}}\tilde{\sigma}_N - \tilde{l}_N^{\text{sf}}\sigma_N}{\tilde{l}_N^{\text{sf}}\tilde{\sigma}_N + \tilde{l}_N^{\text{sf}}\sigma_N} \tilde{B}_N, \quad (\text{A13})$$

$$\tilde{B}_N = -\frac{\sigma_{+,F} - \sigma_{-,F}}{\sigma_{+,F} + \sigma_{-,F}}ej_{\text{app}} \left[\left(\frac{2}{\tilde{l}_F^{\text{sf}}} \frac{\sigma_{+,F}\sigma_{-,F}}{\sigma_{+,F} + \sigma_{-,F}} + \frac{\tilde{\sigma}_N}{\tilde{l}_N^{\text{sf}}} \right) - e^{-2d/\tilde{l}_N^{\text{sf}}} \frac{\tilde{l}_N^{\text{sf}}\tilde{\sigma}_N - \tilde{l}_N^{\text{sf}}\sigma_N}{\tilde{l}_N^{\text{sf}}\tilde{\sigma}_N + \tilde{l}_N^{\text{sf}}\sigma_N} \left(\frac{2}{\tilde{l}_F^{\text{sf}}} \frac{\sigma_{+,F}\sigma_{-,F}}{\sigma_{+,F} + \sigma_{-,F}} - \frac{\tilde{\sigma}_N}{\tilde{l}_N^{\text{sf}}} \right) \right]^{-1}. \quad (\text{A14})$$

D_F and D_N are determined by Eq. (A9). Thus, we determine all coefficients of Eqs. (A7) and (A6).

Appendix B: Scaling procedure of the experimental data

Due to inevitable sample-to-sample variations, and potential changes in interface resistance, limited information can be extracted from changes in the absolute magnitude of α_{eff} on annealing. This however does not preclude an analysis of the changes to Kondo depolarization, provided a method is established to appropriately scale $\alpha_{\text{eff}}(T)$. In this section we will outline the procedure applied to reach the scaled data of Fig. 3(a).

The Kondo expression of Eq. (10) is valid only over a narrow range about $T \approx T_K$, and one of the major successes of the G-G formalism was to accurately describe the evolution of $\rho_K(T)$ from low ($T \ll T_K$) to high T ($T \gg T_K$). It is worth noting at this point the limiting values of ρ_K in these two regimes. At low T the Kondo effect saturates towards a constant scattering rate as the unitary limit is reached, which Kondo proposed to give $\rho_K \rightarrow \rho_m [1 + 2N_0 J \ln(U/T^*)]$. At high T the effect is negligible and ρ_K tends to the classical constant expression for spin-flip scattering via exchange with the ferromagnetic impurity $\rho_K \rightarrow \rho_m$ (i.e. the Korringa rate). Although we cannot *a priori* determine the magnitude in these two regimes for our experimental data, by considering the data of reference [31] and the transition temperatures between the three regimes we can deduce $\rho_K(T \ll T_K)/\rho_K(T \gg T_K) = 1 + 2N_0 J \ln(U/T^*) \approx 1.8$.

Using Eq. (9) we may obtain an experimental estimate of $\rho_K(T)$ at each T_A , which is explicitly dependent upon $\alpha_{\text{eff}}(T)$. (It is worth noting that $\rho_N, \rho_F, l_F^{sf}, l_N^{sf}$ and α_{FM} are measured directly, while η, d and τ_N^{sf}/τ_N are all T -independent, leav-

ing only the scaled value of α_{eff} undetermined in Eq. (9).) To appropriately normalize $\alpha_{\text{eff}}(T)$ to $\alpha_{\text{FM}}(T)$ we linearly scale $\alpha_{\text{eff}}(T)$ (and consequently modify $\rho_K(T)$) in order to reach the correct ratio of ρ_K at low- and high- T (i.e. $\rho_K(T \ll T_K)/\rho_K(T \gg T_K) = 1.8$). Note, that since scaling α_{eff} in this way only ensures the correct ratio of Kondo to classical scattering, we may still fit $1 - \kappa$ to obtain the *magnitude* of the scattering (both Kondo and classical) and therefore deduce η .

Once we have established correct normalization for a single dataset (in this case the unannealed data), we may normalize the remaining data by observing the following relationship from Eq. 10 and Eq. 11:

$$\frac{\rho_K^i(T)d^i}{\rho_K^j(T)d^j} = \frac{d^i \hat{C}_{\text{Fe}}^i}{d^j \hat{C}_{\text{Fe}}^j} = \text{const.} \quad (\text{B1})$$

Where, the superscript i, j denotes values for different T_A . Note this relationship exploits the fact that annealing only serves to increase the magnitude of d and C_{Fe} in ρ_K , both of which are T independent, thus the functional form of $\rho_K(T)$ is independent of T_A . From Eq. (9):

$$\rho_K^i(T)d^i = (1 - r^i \kappa^j) \left[\frac{2\eta(\tau_N^{sf}/\tau_N)}{\rho_N^i} \left(\frac{1}{\lambda_N^i} - \frac{1}{\lambda_N^i + \frac{\lambda_F \rho_F^i}{(1-\alpha^2)\rho_N^i}} \right) \right]^{-1}. \quad (\text{B2})$$

Here r^i is the scaling factor for $\alpha_{\text{eff}}(T)$. Since the ratio of $\rho_K d$ is constant, we minimize the standard deviation of expression B1 by varying r^i , to ensure $\alpha_{\text{eff}}(T)$ is correctly scaled at each T_A .

Appendix C: Mathematical details for the derivation

1. Legendre decomposition of the Boltzmann equation

We first expand the first term in Eq. (12) by the Legendre polynomial. The Bonnet recursion formula is useful to do this.

$$xP_n(x) = \frac{n+1}{2n+1}P_{n+1}(x) + \frac{n}{2n+1}P_{n-1}(x), \quad (\text{C1})$$

for $n \geq 1$.

$$\begin{aligned} \frac{\hbar k_z}{m} \partial_z g_s(z, \mathbf{k}) &= \frac{\hbar k}{m} \sum_{n=0}^{\infty} \partial_z g_s^{(n)}(z, k) \cos \theta_{\mathbf{k}} P_n(\cos \theta_{\mathbf{k}}) \\ &= \frac{\hbar k}{m} \partial_z g_s^{(0)}(z, k) P_1(\cos \theta_{\mathbf{k}}) + \frac{\hbar k}{m} \sum_{n=1}^{\infty} \partial_z g_s^{(n)}(z, k) \left[\frac{n+1}{2n+1} P_{n+1}(\cos \theta_{\mathbf{k}}) + \frac{n}{2n+1} P_{n-1}(\cos \theta_{\mathbf{k}}) \right] \\ &= \frac{\hbar k}{m} \sum_{n=0}^{\infty} \left[\frac{n}{2n-1} \partial_z g_s^{(n-1)}(z, k) + \frac{n+1}{2n+3} \partial_z g_s^{(n+1)}(z, k) \right] P_n(\cos \theta_{\mathbf{k}}). \end{aligned} \quad (\text{C2})$$

The second term in Eq. (12) is

$$-\frac{eE}{\hbar} \partial_{k_z} f_s^0(k) = -\frac{eE}{\hbar} \partial_k f_s^0(k) \times P_1(\cos \theta_{\mathbf{k}}). \quad (\text{C3})$$

The right-hand side of Eq. (12) is

$$-\frac{g_s(z, \mathbf{k}) - \overline{g_s(z, k)}}{\tau_s^{\text{sc}}(z, k)} - \frac{g_s(\mathbf{k}) - \overline{g_s(z, k)}}{\tau_s^{\text{sf}}(z, k)} = -\frac{g_s^{(0)}(z, k) - g_{-s}^{(0)}(z, k)}{\tau_s^{\text{sf}}(z, k)} - \frac{1}{\tau_s(z, k)} \sum_{n=1}^{\infty} g_s^{(n)} P_n(\cos \theta_{\mathbf{k}}), \quad (\text{C4})$$

where $1/\tau_s = 1/\tau_s^{\text{sc}} + 1/\tau_s^{\text{sf}}$.

In summary, Eq. (12) is equivalent to

$$\begin{aligned} & \sum_{n=0}^{\infty} \left[\frac{\hbar k}{m} \frac{n}{2n-1} \partial_z g_s^{(n-1)}(z, k) + \frac{\hbar k}{m} \frac{n+1}{2n+3} \partial_z g_s^{(n+1)}(z, k) - \frac{eE}{\hbar} \partial_k f_s^0(k) \delta_{n,1} \right] P_n(\cos \theta_{\mathbf{k}}) \\ &= -\frac{g_s^{(0)}(z, k) - g_{-s}^{(0)}(z, k)}{\tau_s^{\text{sf}}(z, k)} - \frac{1}{\tau_s(z, k)} \sum_{n=1}^{\infty} g_s^{(n)} P_n(\cos \theta_{\mathbf{k}}). \end{aligned} \quad (\text{C5})$$

Since $\{P_n\}$ forms an orthogonal set of polynomials, each coefficient should satisfy the equation. In Ref. [17], if the spin diffusion length is much larger than the mean free path of conduction electrons, $g_s^{(2)}$ (and higher order terms) can be neglected. The coefficients of P_0 and P_1 gives Eqs. (13) and (14), once $g_s^{(2)}$ is neglected.

2. Sommerfeld expansion formula

In this section, we present the Sommerfeld formula for low temperature. In the main text, we keep terms up to $O(T)$, we here present the formula up to $O(T^3)$ for more motivated readers.

The Sommerfeld expansion formula for a differentiable function H is

$$\int \frac{H(\varepsilon)}{e^{\beta(\varepsilon-\mu)} + 1} d\varepsilon = \int^{\mu} H(\varepsilon) d\varepsilon + \frac{\pi^2}{6\beta^2} H'(\mu) + O(T^4). \quad (\text{C6})$$

In a compact form, as far as quantities after integration over ε is concerned,

$$\frac{1}{e^{\beta(\varepsilon-\mu)} + 1} = \Theta(\mu - \varepsilon) + \frac{\pi^2}{6\beta^2} \delta'(\mu - \varepsilon) + O(T^4). \quad (\text{C7})$$

When a transport property is concerned, it is convenient to take the derivative with respect to ε .

$$\frac{\partial}{\partial \varepsilon} \frac{1}{e^{\beta(\varepsilon-\mu)} + 1} = -\delta(\mu - \varepsilon) - \frac{\pi^2}{6\beta^2} \delta''(\mu - \varepsilon) + O(T^4). \quad (\text{C8})$$

3. Integrals including the Kondo scattering rate

In this section, we perform the following integration for a general $G(\varepsilon)$

$$\int d\varepsilon G(\varepsilon) \gamma(\varepsilon) \partial_{\varepsilon} f^0, \quad (\text{C9})$$

where $f^0 = [1 + e^{\beta(\varepsilon-\mu^0)}]^{-1}$ and $\gamma(\varepsilon)$ is defined by Eq. (24). Here and from now on, we denote $k = \sqrt{2m\varepsilon}/\hbar$, $k' = \sqrt{2m\varepsilon'}/\hbar$ and so on, appearing in integrations with respect to ε and ε' . Also, we define $k^F = \sqrt{2m\mu^0}/\hbar$ which is the Fermi wave vector. We generalize the approach taken by Kondo [31] here.

First we perform the summation in $\gamma(\varepsilon)$.

$$\begin{aligned} \gamma(\varepsilon) &= \frac{1}{V} \sum_{\mathbf{k}} \frac{f^0(\mathbf{k})}{\varepsilon(\mathbf{k}) - \varepsilon} = \frac{1}{8\pi^3} \int d^3 k' \frac{f^0(\mathbf{k}')}{\varepsilon(\mathbf{k}') - \varepsilon} = \frac{1}{2\pi^2} \int dk' k'^2 \frac{f^0(k')}{\varepsilon(k') - \varepsilon} \\ &= \frac{m}{\pi^2 \hbar^2} \int dk' \frac{\varepsilon(k')}{\varepsilon(k') - \varepsilon} f^0(k') = \frac{m}{\pi^2 \hbar^2} \int dk' f^0(k') + \frac{m\varepsilon}{\pi^2 \hbar^2} \int dk' \frac{f^0(k')}{\varepsilon(k') - \varepsilon}. \end{aligned} \quad (\text{C10})$$

The first integral can be given by the Sommerfeld expansion Eq. (C7).

$$\frac{m}{\pi^2 \hbar^2} \int dk' f^0(k') = \frac{m \sqrt{m}}{\pi^2 \hbar^3} \int d\varepsilon' \frac{1}{\sqrt{2\varepsilon'}} f^0(\varepsilon') = 2N_0. \quad (\text{C11})$$

To perform the second integral,

$$\frac{m\varepsilon}{\pi^2\hbar^2} \int dk' \frac{f^0(k')}{\varepsilon(k') - \varepsilon} = \frac{2m^2\varepsilon}{\pi^2\hbar^4} \int dk' \frac{f^0(k')}{k'^2 - k^2} = -\frac{m^2\varepsilon}{\pi^2\hbar^4 k} \int dk' \ln \left| \frac{k - k'}{k + k'} \right| \partial_{k'} f^0 = -\frac{mk}{2\pi^2\hbar^2} \int d\varepsilon' \ln \left| \frac{k - k'}{k + k'} \right| \partial_{\varepsilon'} f^0. \quad (\text{C12})$$

We are now ready to perform the integral in Eq. (C9).

$$\begin{aligned} \int d\varepsilon G(\varepsilon) \gamma(\varepsilon) \partial_{\varepsilon} f^0 &= 2N_0 \int d\varepsilon G(\varepsilon) \partial_{\varepsilon} f^0 - \frac{m}{2\pi^2\hbar^2} \int d\varepsilon d\varepsilon' k G(\varepsilon) \ln \left| \frac{k - k'}{k + k'} \right| \partial_{\varepsilon} f^0 \partial_{\varepsilon'} f^0 \\ &= 2N_0 \int d\varepsilon G(\varepsilon) \partial_{\varepsilon} f^0 - \frac{m\sqrt{m}}{\sqrt{2}\pi^2\hbar^3} \int d\varepsilon d\varepsilon' \sqrt{\varepsilon} G(\varepsilon) \ln \left| \frac{\sqrt{\varepsilon} - \sqrt{\varepsilon'}}{\sqrt{\varepsilon} + \sqrt{\varepsilon'}} \right| \partial_{\varepsilon} \frac{1}{1 + e^{\beta(\varepsilon - \mu^0)}} \partial_{\varepsilon'} \frac{1}{1 + e^{\beta(\varepsilon' - \mu^0)}}. \end{aligned} \quad (\text{C13})$$

The first integral is given by the Sommerfeld expansion Eq. (C8). For the second term, exact substitution of $\varepsilon = \varepsilon'$ yields divergence, however, we may still calculate the temperature dependence of the term. By substituting $X = \beta(\varepsilon - \mu^0)$ and $X' = \beta(\varepsilon' - \mu^0)$, the second term is

$$\begin{aligned} & -\frac{m\sqrt{m}}{\sqrt{2}\pi^2\hbar^3} \int d\varepsilon d\varepsilon' \sqrt{\varepsilon} G(\varepsilon) \ln \left| \frac{\sqrt{\varepsilon} - \sqrt{\varepsilon'}}{\sqrt{\varepsilon} + \sqrt{\varepsilon'}} \right| \partial_{\varepsilon} \frac{1}{1 + e^{\beta(\varepsilon - \mu^0)}} \partial_{\varepsilon'} \frac{1}{1 + e^{\beta(\varepsilon' - \mu^0)}} \\ &= -\frac{m\sqrt{m}}{\sqrt{2}\pi^2\hbar^3} \int dXdX' \sqrt{k_B T X + \mu^0} G(k_B T X + \mu^0) \ln \left| \frac{\sqrt{k_B T X + \mu^0} - \sqrt{k_B T X' + \mu^0}}{\sqrt{k_B T X + \mu^0} + \sqrt{k_B T X' + \mu^0}} \right| \partial_X \frac{1}{1 + e^X} \partial_{X'} \frac{1}{1 + e^{X'}} \\ &\approx -\frac{m\sqrt{m}}{\sqrt{2}\pi^2\hbar^3} \sqrt{\mu^0} G(\mu^0) \int dXdX' \left(\ln \frac{k_B T}{\mu^0} + \ln \frac{X - X'}{4} \right) \partial_X \frac{1}{1 + e^X} \partial_{X'} \frac{1}{1 + e^{X'}}. \end{aligned} \quad (\text{C14})$$

Here we expanded with respect to $k_B T$, which is small compared to μ^0 . We drop the second contribution $\ln(X - X')/4$ since it gives a much smaller contribution than the $\ln k_B T / \mu^0$ contribution at low temperature. Thus we keep only the logarithmic term.

$$-\frac{m\sqrt{m}}{\sqrt{2}\pi^2\hbar^3} \int d\varepsilon d\varepsilon' \sqrt{\varepsilon} G(\varepsilon) \ln \left| \frac{\sqrt{\varepsilon} - \sqrt{\varepsilon'}}{\sqrt{\varepsilon} + \sqrt{\varepsilon'}} \right| \partial_{\varepsilon} \frac{1}{1 + e^{\beta(\varepsilon - \mu^0)}} \partial_{\varepsilon'} \frac{1}{1 + e^{\beta(\varepsilon' - \mu^0)}} \approx -\frac{m}{2\pi^2\hbar^2} k^F G(\mu^0) \ln \frac{k_B T}{\mu^0}. \quad (\text{C15})$$

By using $N_0 = mk^F / 2\pi^2\hbar^2$, for low temperature, the following replacement is valid under an energy integration.

$$\gamma(\varepsilon) \partial_{\varepsilon} f^0 \rightarrow -N_0 \ln \frac{k_B T}{\mu^0} \delta(\varepsilon - \mu^0), \quad (\text{C16})$$

which is Eq. (26).

Appendix D: Error analysis

Errors in the parameters I_N^{sf} , α_{eff} , ρ_N , ρ_F (experimentally determined) and κ (determined through a normalization procedure), as well as uncertainty from our fitting method are our main concern in establishing uncertainty in the extracted values of η . All other parameters are constrained through the previous work, and potential errors in such quantities are not considered.

Although both ρ_N and ρ_F are measured directly from R_N and R_F , with very small random noise, these quantities suffer from experimental uncertainty in the wire cross sectional area (through $\rho = RA/L$), particularly a non-rectangular shape. This uncertainty is random between devices, but *systematic* across all T within a single device. It is estimated at the level of $\approx 5\%$ from SEM images of the wire edge profile. In our measurement of ΔR_{NL} we observe a baseline noise floor of around 1 nV (at a modulation frequency of 13 Hz). This corresponds to $\approx 3 \mu\Omega$ in our measurements and is an absolute

noise source independent of signal size. When fitting $\Delta R_{\text{NL}}(L)$ at each temperature to extract I_N^{sf} and α_{eff} , the uncertainties in ρ and ΔR_{NL} are used to weight a least means square minimization fit, with estimated errors in I_N^{sf} and α_{eff} arising from combining these errors with the fitting residuals. In reality, the parameters I_N^{sf} and α_{eff} , are limited in precision by the relative uncertainty in the cross-sectional area measurements, and are therefore largely independent of T . Although we may obtain α_{eff} and I_N^{sf} by fitting $\Delta R_{\text{NL}}(L, T)$, using a literature value of I_N^{sf} as a constraint on the signal magnitude, the magnitude of α_{eff} is poorly constrained, due to the inherent difficulty in precisely measuring the ferromagnet/normal metal interface resistance. Consequently, errors that are independent of T dominate the extracted values of α_{eff} .

With the errors for I_N^{sf} and α_{eff} established, it remains to estimate the uncertainty in κ , before determining η . As both α_{FM} and α_{eff} are broadly of a similar magnitude, and $\kappa = \alpha_{\text{eff}} / \alpha_{\text{FM}}$, the systematic errors in each quantity could, in principle, give an error larger than the estimated value of $1 - \kappa$ (typically $1 - \kappa$

is around 10 %, while errors in α are around 5 % to 10 %). However, this systematic error is irrelevant for the normalization procedure we use, and is one of the key advantages to our method: As any error in α are largely T -independent (errors from both fitting and estimates of interface resistance), they make no impact on the overall normalization factor [r in Eq. (B2)], since any systematic error in α_{eff} or α_{FM} is intrinsically compensated by r . Thus we can estimate the error in κ solely from the uncertainty in the normalization procedure. To obtain this value, we realize that the process of minimizing the standard deviation of $\rho_K^1 d^1 / \rho_K^i d^i$ (our normalization procedure) is identical to a linear regression of $y^i = A^i(1 - r^i \kappa^i)$, where $y^i = \rho_K^1 / \rho_N^i [1/l_N^{\text{sf},i} - 1/(l_N^{\text{sf},i} + \rho_N^i l_F^{\text{sf}}[(1 - \alpha_{\text{FM}}^2)\rho_F]^{-1})]$, and $A^i = C_{\text{Fe}}^1 d^1 \tau_N^{\text{sf}} / 2C_{\text{Fe}}^i d^i \eta \tau_N$. Once again, we use the superscript i to denotes a given dataset (i.e. a given T_A), with $i = 1$ representing the unannealed data (which can be scaled exactly, see Appendix B). As we can establish both y and κ for a given T_A , we may therefore estimate the uncertainty in r , and so the relative error in our scaled κ , from the residuals of a least mean squares fit of $y = A(1 - r\kappa)$ with A and r as free variables. These estimates are the error bars shown in Fig 4.

The final challenge is to incorporate all these errors together for our final fitting procedure to estimate η . η is found from fitting $\eta\rho_K$ from experimental data using our models for ρ_K , i.e. through rearranging Eq. (9). Through the discussed procedure we now have estimates for all parameters in this equation, including errors for the experimentally determined quantities (κ , ρ_N , ρ_F , α , l_N^{sf}). To establish the error on the experimental $\eta\rho_K$ we use Monte Carlo sampling assuming Gaussian distributed uncertainties for all quantities (*via* the NIST uncertainty machine [46]) to account for the combination of all uncertainties in Eq. (9). The extracted errors are subsequently used as weightings for fitting $\eta\rho_K$ to either the G-G or Kondo model, again using a least-mean-square approach. The extracted parameter uncertainties are shown in Fig. 4 for each method, with the final errors for η in panel (c). All quoted errors are a single standard deviation, including those shown for the extracted values of T_K . Most errors are relative rather than absolute, and so data at large T_A appear more error prone than those at low T_A , despite the larger κ . In calculating $\hat{\eta}$ an unweighted average is taken, with the uncertainty in this case quoted as the standard deviation in the eight values.

-
- [1] J. Bass and W. P. Pratt, J. Phys. Condens. Matter **19**, 183201 (2007).
 - [2] K. Uchida, S. Takanashi, K. Harii, J. Ieda, W. Koshibae, K. Ando, S. Maekawa, and E. Saitoh Nature **455**, 778 (2008).
 - [3] S. O. Valenzuela and M. Tinkham, Nature **442**, 176 (2006).
 - [4] Y. Tserkovnyak, A. Brataas, G. E. W. Bauer, and B. Halperin, Rev. Mod. Phys. **77**, 1375 (2005).
 - [5] M. Johnson and R. H. Silsbee, Phys. Rev. Lett. **55**, 1790 (1985).
 - [6] F. J. Jedema, A. T. Filip, and B. J. van Wees, Nature **410**, 345 (2001).
 - [7] S. Garzon, I. Žutić, and R. A. Webb, Phys. Rev. Lett. **94**, 176601 (2005).
 - [8] R. Godfrey and M. Johnson, Phys. Rev. Lett. **96**, 136601 (2006).
 - [9] T. Kimura and Y. Otani, Phys. Rev. Lett. **99**, 196604 (2007).
 - [10] T. Yang, T. Kimura, and Y. Otani, Nat. Phys. **4**, 851 (2008).
 - [11] F. L. Bakker, A. Slachter, J.-P. Adam, and B. J. van Wees, Phys. Rev. Lett. **105**, 136601 (2010).
 - [12] G. Mihajlović, J. E. Pearson, S. D. Bader, and A. Hoffmann, Phys. Rev. Lett. **104**, 237202 (2010).
 - [13] A. Slachter, F. L. Bakker, J.-P. Adam, and B. van Wees, Nat. Phys. **6**, 879 (2010).
 - [14] Y. Fukuma, L. Wang, H. Idzuchi, S. Takahashi, S. Maekawa, and Y. Otani, Nat. Mater. **10**, 527 (2011).
 - [15] Y. Niimi, D. Wei, H. Idzuchi, T. Wakamura, T. Kato, and Y. C. Otani, Phys. Rev. Lett. **110**, 016805 (2013).
 - [16] M. A. M. Gijs and G. E. W. Bauer, Adv. Phys. **46**, 285 (1997).
 - [17] T. Valet and A. Fert, Phys. Rev. B **48**, 7099 (1993).
 - [18] R. J. Elliott, Phys. Rev. **96**, 266 (1954).
 - [19] V. V. Andreev and V. I. Gerasimenko, Sov. Phys. JETP **35**, 846 (1959).
 - [20] Y. Yafet, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic, New York, 1963), Vol. 14.
 - [21] T. Kimura, T. Sato, and Y. Otani, Phys. Rev. Lett. **100**, 066602 (2008).
 - [22] F. Casanova, A. Sharoni, M. Erekhinsky, and I. K. Schuller, Phys. Rev. B **79**, 184415 (2009).
 - [23] Y. Otani and T. Kimura, Philos. Trans. A. Math. Phys. Eng. Sci. **369**, 3136 (2011).
 - [24] M. Erekhinsky, F. Casanova, I. K. Schuller, and A. Sharoni, Appl. Phys. Lett. **100**, 212401 (2012).
 - [25] T. Kimura, N. Hashimoto, S. Yamada, M. Miyao, and K. Hamaya, NPG Asia Mater. **4**, e9 (2012).
 - [26] H. Zou and Y. Ji, Appl. Phys. Lett. **101**, 082401 (2012).
 - [27] E. Villamor, M. Isasa, L. E. Hueso, and F. Casanova, Phys. Rev. B **87**, 094417 (2013).
 - [28] J. T. Batley, M. C. Rosamond, M. Ali, E. H. Linfield, G. Burnell, and B. J. Hickey, Phys. Rev. B **92**, 220420(R) (2015).
 - [29] L. O'Brien, M. J. Erickson, D. Spivak, H. Ambaye, R. J. Goyette, V. Lauter, P. A. Crowell, and C. Leighton, Nat. Commun. **5**, 3927 (2014).
 - [30] L. O'Brien, D. Spivak, J. S. Jeong, K. A. Mkhoyan, P. A. Crowell, and C. Leighton, Phys. Rev. B **93**, 014413 (2016).
 - [31] J. Kondo, Prog. Theor. Phys. **32**, 37 (1964).
 - [32] C. Domenicali and E. Christenson, J. Appl. Phys. **32**, 2450 (1961).
 - [33] K. Hamaya, T. Kurokawa, S. Oki, S. Yamada, T. Kanashima, and T. Taniyama, arXiv:1603.08599.
 - [34] T. Taniyama, N. Fujiwara, Y. Kitamoto, and Y. Yamazaki, Phys. Rev. Lett., **90**, 016601 (2003).
 - [35] A. Fert, J.-L. Duvail, and T. Valet, Phys. Rev. B **52**, 6513 (1995).
 - [36] D. Goldhaber-Gordon, J. Göres, and M. A. Kastner, H. Shtrikman, D. Mahalu, and U. Meirav, Phys. Rev. Lett. **81**, 5225 (1998).
 - [37] W. G. van der Wiel, S. De Franceschi, T. Fujisawa, J. M. Elzerman, S. Tarucha, and L. P. Kouwenhoven, Science **289**, 2105 (2000).
 - [38] J. J. Parks, A. R. Champagne, G. R. Hutchison, S. Flores-Torres, H. D. Abruna, and D. C. Ralph, Phys. Rev. Lett. **99**, 026601 (2007).
 - [39] S. Takahashi and S. Maekawa, Phys. Rev. B **67** 052409 (2003).

- [40] T. Hutchison and J. Reekie, Phys. Rev. **83**, 854 (1951).
- [41] J. Mydosh, in *Spin Glasses: an Experimental Introduction*, (Taylor and Francis, London, 1993).
- [42] F. Beuneu and P. Monod, Phys. Rev. B **18**, 2422 (1978).
- [43] N. W. Ashcroft and N. D. Mermin, in *Solid State Physics*, (Holt, Rinehart, and Winston, 1976).
- [44] P. Monod, Phys. Rev. Lett. **19**, 1113 (1967).
- [45] H. Idzuchi, Y. Fukuma, L. Wang and Y. Otani, App. Phys. Lett. **101**, 022415 (2012).
- [46] T. Lafarge and A. Possolo, NCLSI Measure Journal of Measurement Science **10**, 20 (2015); see <http://uncertainty.nist.gov>
- [47] $\mu\text{mol/mol}$ is equivalent to ‘parts per million’
- [48] We emphasize that each region is specified by subscripts. Note that the *superscript* F that appears in k_s^F for instance refers to at the Fermi surface.
- [49] One can verify explicitly that κ has the same value for both definitions.
- [50] $\int_{x(C_{\text{Fe}}=2600\mu\text{mol/mol})}^{\infty} [1 - \text{erf}(x/\lambda_{\text{Fe}})] dx = 560\mu\text{mol/mol} \cdot \lambda_{\text{Fe}}$
- [51] See Eq. (12) of Ref. [31].